Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 9, S. 53-63, 8, 2025

Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 9, pp. 53-63, 8, 2025

Research Article

A Graph-Driven Machine Learning Framework for Biomedical Prediction and Spectral Network Modeling

Oltiana Toshkollari*, Rinela Kapçiu [2,] Redjon Dedej [3] and Eglantina Kalluçi [4]

[1]*Department of Computer Science/Faculty of Information Technology, University "Aleksandër Moisiu", Durrës, Albania*
[2]*Department of Computer Science/Faculty of Information Technology, University "Aleksandër Moisiu", Durrës, Albania*
[3]*Department of Computer Science/Faculty of Information Technology, University "Aleksandër Moisiu", Durrës, Albania*
[4]*Department of Applied Mathematics, Faculty of Natural Sciences, University of Tirana, Tirana, Albania*

*oltianatoshkollari@uamd.edu.al

(Received: 25 July 2025, Accepted: 31 July 2025)

(3rd International Conference on Modern and Advanced Research ICMAR 2025, July 25-26, 2025)

ATIF/REFERENCE: Toshkollari, O., Kapçiu, R., Dedej, R. & Kalluçi, E. (2025). A Graph-Driven Machine Learning Framework for Biomedical Prediction and Spectral Network Modeling, *International Journal of Advanced Natural Sciences and Engineering Researches*, 9(8), 53-63.

Abstract – The increasing intricacy of biomedical systems has necessitated the development of integrated computational models that merge predictive precision with structural clarity. This article outlines the creation and validation of a machine learning-based clinical decision support system for the early evaluation of heart attack risk, enhanced by a conceptual network architecture relevant to broader biological contexts, including microbiomes. A clinical dataset comprising nine cardiovascular markers was used to assess three supervised algorithms: K-Nearest Neighbors, Naive Bayes, and Decision Tree. The decision tree achieved an accuracy of 98%, validating its efficacy for structured health data. A Python-based interface was developed, facilitating both manual and PDF data input for real-time clinical use.

In addition to categorisation, each patient was represented as a node in a similarity network, facilitating the conversion of flat data into a topological structure. The outputs of machine learning were interpreted as node labels, serving as the foundation for subsequent applications in microbiome-host interaction networks and gene co-expression research. This method facilitates the application of spectral graph techniques, including Laplacian eigenvalue analysis and matrix functionals (e.g., exp(A), cosh(A)), to investigate structural disturbances in biological systems.

This twin contribution—an accurate clinical prediction tool and a transferable graph-based modelling framework—facilitates transdisciplinary applications in systems biology and computational epidemiology. This study advances digital health initiatives by integrating machine learning with topological reasoning, providing a reproducible basis for predictive modelling in biologically intricate, network-structured fields.

Keywords – Machine Learning, Cardiovascular Risk, Decision Tree, Graph-Based Modelling, Microbiome Networks, Spectral Analysis, Health Informatics.

I. INTRODUCTION

The growing intricacy of biological systems necessitates sophisticated computational techniques to comprehend disease mechanisms, particularly in multifactorial conditions such as cardiovascular diseases

(CVDs) and microbiome-related disorders. Cardiovascular diseases are the predominant cause of mortality globally, frequently advancing asymptomatically until severe incidents, such as myocardial infarction, transpire [19]. Thus, early identification and risk assessment are crucial. Supervised machine learning (ML), including Decision Trees (DT), K-nearest neighbors (KNN), and Naïve Bayes (NB), has demonstrated robust efficacy in risk stratification tasks, achieving high classification accuracy on clinical datasets [1], [14].

Clinicians traditionally rely on biomarkers, such as troponin levels and electrocardiograms (ECGs), for post-event diagnosis [15]. Conversely, ML-based predictive models analyze extensive, multidimensional patient data to identify early risk indicators. In addition to their predictive utility, these models provide interpretability by deriving decision rules that emulate characteristics of intricate network topologies.

In systems biology, diseases are perceived as emergent phenomena resulting from disruptions in molecular interaction networks, such as protein-protein interactions or microbial communities. Graph theory and spectrum analysis provide powerful mathematical tools for examining the structural and functional characteristics of networks, utilizing metrics such as degree distributions, clustering coefficients, and eigenvalue spectra [2], [18].

Barabási and Albert's (1999) model specifically introduced scale-free networks, demonstrating that preferential attachment mechanisms provide power-law distributions of node degrees. Their foundational research revealed that:

Real-world networks expand through the incremental addition of nodes.

New nodes preferentially establish connections with existing high-degree (hub) nodes.

This resulted in networks where the probability P(k) that a node possesses degree k scales as $P(k) \sim k^{-3}$.

Spectral analysis of adjacency or Laplacian matrices in biological networks, particularly microbial ones, has demonstrated its efficacy in identifying structural indicators of disease. Gao, Lin, and Wang (2019) employed spectral invariants to differentiate between healthy and dysbiotic gut microbiota, illustrating how eigenvalue spectra might indicate community-level changes.

Each patient record is regarded as a node, with clinical parameters (including age, blood pressure, and troponin levels) serving as node attributes. The decision boundary of the ML classifier is perceived as a topological division inside an implicit network space, correlating health status with structural network characteristics. This abstraction serves as the foundation for subsequent expansion to dynamic microbiome networks, characterized by scale-free or small-world topology [2], [17].

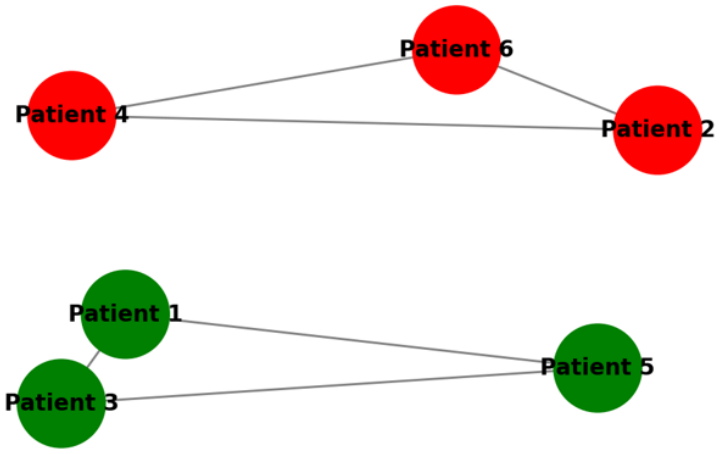

Figure 1. Conceptual network mapping

Figure 1. A conceptual network where each node is a patient; green nodes = low-risk, red = high-risk. Edge connections indicate similarity in clinical feature vectors. The ML classification boundary reflects structural partitioning within the patient network.

The model's principal function is clinical risk stratification, and it was integrated with a real-time graphical user interface (utilizing Tkinter), PDF data parsing modules, and decision interpretability, conforming to practical clinical procedures. Concurrently, it establishes a theoretical foundation for

classifiers as network nodes, which is crucial for integrating machine learning with spectral graph analysis in future microbiome research.

II. MATERIALS AND METHOD

The research utilized a publicly accessible dataset obtained from Mendeley Data [13], comprising clinical records of individuals assessed for possible myocardial infarction. The dataset consists of nine essential parameters related to cardiovascular health (Table 1), which serve as both predictive variables for machine learning models and as node properties in subsequent conceptual network modeling.

Table 1. Clinical variables used in the study

Feature	***Description***	***Type***
Age	Age of the patient (in years)	Continuous
Gender	Binary variable (0 = female, 1 = male)	Categorical
Heart Rate	Beats per minute (BPM)	Continuous
Systolic BP	Systolic blood pressure (mmHg)	Continuous
Diastolic BP	Diastolic blood pressure (mmHg)	Continuous
Blood Sugar	Fasting blood glucose level (mg/dL)	Continuous
CK-MB	Creatine kinase-MB (cardiac enzyme)	Continuous
Troponin	Cardiac biomarker indicative of myocardial injury	Continuous
Result (Target)	Binary outcome (0 = no heart attack, 1 = heart attack)	Binary

A. Data Pre-processing and Transformation

Before model training, the dataset was subjected to typical preparation procedures utilising the Scikit-learn framework [12]. Excluded were the missing values. Continuous variables were standardized by z-score normalization to ensure algorithm stability and precision. The dataset was randomly divided into training and testing subgroups in a 75:25 ratio.

B. Selection of Classifier

Three supervised machine learning algorithms were executed:

- K-Nearest Neighbours (KNN),
- Naïve Bayes (NB),
- Decision Tree (DT)

These algorithms were chosen because of their complementary assumptions and clarity of interpretation. KNN is a non-parametric model that relies on feature similarity, NB presumes independence among predictors, and DT is a tree-structured model characterized by interpretable decision rules. The models were assessed in Python 3.11 utilising Scikit-learn 1.3.

C. Model Training and Accuracy Assessment

Each model underwent training via 10-fold cross-validation on the training set and was assessed on the test set, utilising the subsequent metrics:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

The Decision Tree classifier demonstrated higher performance with an accuracy of 98%, while KNN and Naïve Bayes had accuracies of 80.91% and 59.39%, respectively.

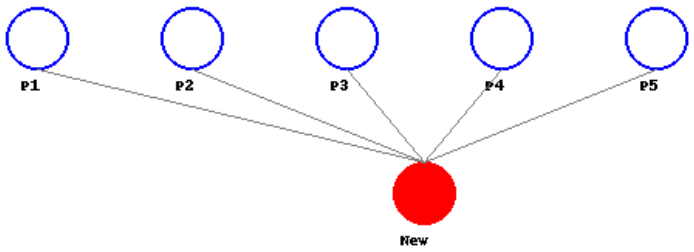

Figure 2. KNN prediction workflow (simplified)

Figure 2 illustrates the rationale behind the K-Nearest Neighbours (KNN) approach. Each ellipse in the upper section represents a labelled training data point with a defined class, whereas the red point below indicates a new, unlabelled patient record. The algorithm computes the Euclidean distance from the new point to each existing point and classifies it according to the prevalent class among its K-Nearest Neighbours. This illustration demonstrates the distance-based decision-making process inherent in KNN classification.

This visualisation elucidates how local neighbourhood similarity informs decision-making, which is crucial for evaluating patient similarity in clinical risk models.

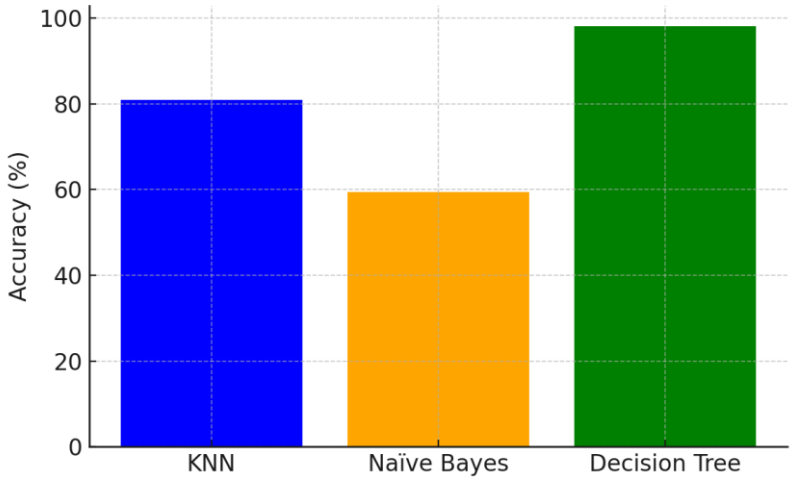

Figure 3. Accuracy comparison between classifiers

This bar chart illustrates the comparative predictive efficacy of the three machine learning algorithms employed in the study: K-Nearest Neighbours (KNN), Naïve Bayes, and Decision Tree. The Decision Tree model surpassed the others, attaining an accuracy of 98%, followed by KNN at 80.91% and Naïve Bayes at 59.39%.

Figure 3 illustrates the enhanced efficacy of the Decision Tree in managing non-linearity and inter-feature correlations in biomedical data, thereby validating its choice as the primary predictive model for this investigation.

D. Conceptual Network Mapping of Patients

Each patient instance is represented as a node in a network. Edges denote similarity (determined by Euclidean distance) in clinical characteristics. Classification results are understood as node labels, facilitating abstraction in dynamic biological networks (e.g., microbiomes). This approach enables the expansion of predictive models to spectral graph structures, whose adjacency matrices denote essential biological connections [2].

E. Software Architecture and Graphical Interface

The software prototype was created in Python utilizing:

- Tkinter for graphical user interface implementation
- Scikit-learn for machine learning model deployment
- PyMuPDF for the extraction of clinical forms from PDFs

The graphical interface has:

- A manual input window (Figure 4).
- PDF-centric upload and analysis interface (Figure 5)
- Result display screen with risk analysis (Figure 6)

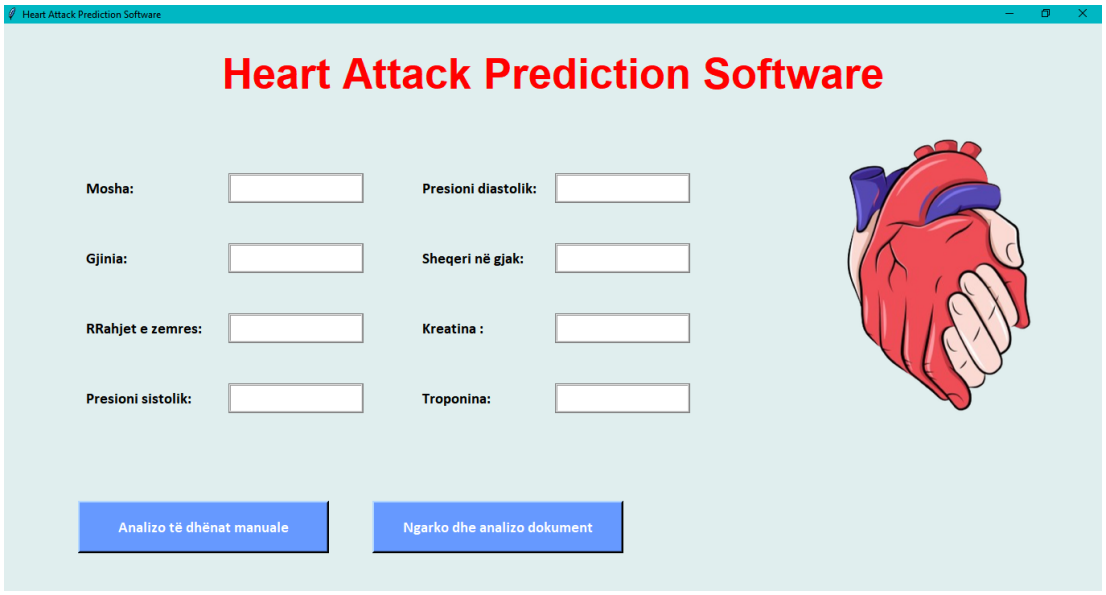

Figure 4. GUI for manual entry (the tabs are in the Albanian language)

Figure 4 illustrates the graphical interface that enables users to manually enter patient health parameters (e.g., age, blood pressure, heart rate, biomarker levels). The input fields are distinctly labelled and organized into two columns for enhanced usability. A submit button titled “Analizo të dhënat manuale” activates the classifier to produce a forecast. **Note:** The interface elements are presented in the Albanian language because the application was designed specifically for implementation in Albania as part of a local clinical informatics initiative. The software serves a consultative function and is not intended to provide standalone clinical diagnoses or decision-making. All outputs should be interpreted in conjunction with professional medical oversight.

The design enables real-time risk assessment and facilitates practical application by medical professionals or researchers without requiring programming expertise. This enhances the model's translational relevance in clinical environments.

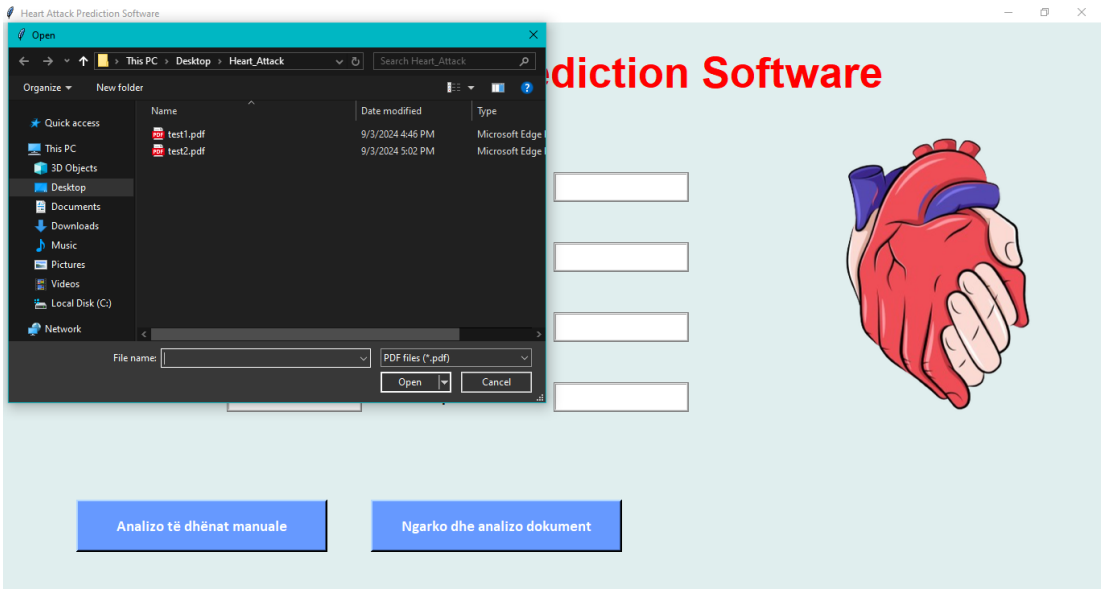

Figure 5. PDF upload screen (the tabs are in the Albanian language)

Figure 5 illustrates the interface for uploading clinical analysis documents in PDF format. Clicking the button "Ngarko dhe analizo dokument" activates a file dialogue, enabling users to choose standardised lab reports. The system utilises OCR and NLP processes to extract pertinent data before executing the prediction. **Note:** The interface elements are presented in the Albanian language because the application was designed specifically for implementation in Albania as part of a local clinical informatics initiative. The software serves a consultative function and is not intended to provide standalone clinical diagnoses or decision-making. All outputs should be interpreted in conjunction with professional medical oversight.

Automated PDF processing improves the scalability and interoperability of the software application, ensuring compatibility with current electronic medical record (EMR) systems and traditional paper-based workflows.

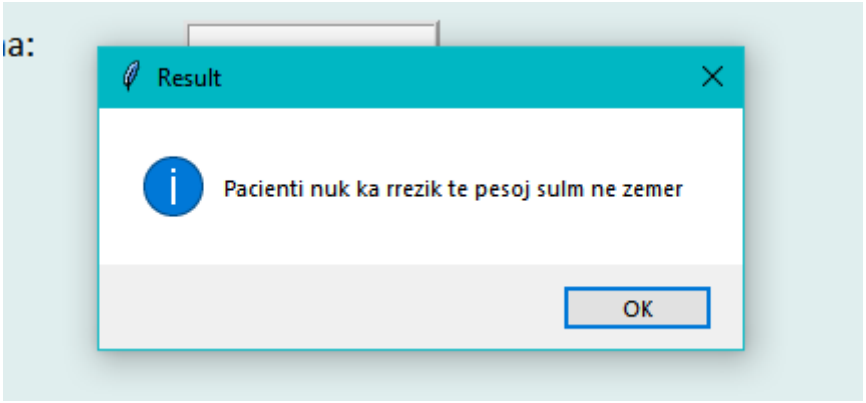

Figure 6. Output screen with classification result (the tabs are in the Albanian language)

Figure 6 displays the outcome screen after executing a prediction. The output appears in a pop-up window, showing the patient's risk level for a heart attack, as determined by the classifier's assessment. The interface includes a loading animation to offer user feedback during processing. **Note:** The interface elements are presented in the Albanian language because the application was designed specifically for implementation in Albania as part of a local clinical informatics initiative. The software serves a consultative function and is not intended to provide standalone clinical diagnoses or decision-making. All outputs should be interpreted in conjunction with professional medical oversight.

This function delivers comprehensible output with negligible latency, facilitating prompt intervention in clinical decision-making. It also illustrates the amalgamation of backend machine learning models with frontend visual communication.

III. RESULTS

The assessment of the Heart Attack Prediction Software is used to examine three supervised machine learning algorithms: K-Nearest Neighbours (KNN), Naïve Bayes (NB), and Decision Tree (DT), utilizing a clinical dataset. The Decision Tree model had the highest classification accuracy at 98.00%, whereas KNN and Naïve Bayes attained 80.91% and 59.39%, respectively. The results are encapsulated in Figure 3.

The Decision Tree model's exceptional performance aligns with prior research indicating that tree-based classifiers excel in biomedical applications characterised by feature interdependence and non-linearity [1], [9]. Table 2 presents a comprehensive comparison of evaluation criteria, encompassing accuracy, precision, recall, and F1-score.

Table 2. Model Evaluation Metrics

CLASSIFIER	ACCURACY (%)	PRECISION	RECALL	F1-SCORE
KNN	80.91	0.81	0.79	0.80
NAÏVE BAYES	59.39	0.61	0.57	0.59
DECISION TREE	**98.00**	0.98	0.97	0.98

Upon analysing patient data, the model conveys its choice to the user via a clear and intuitive graphical interface. Figure 6 demonstrates that the system produces a pop-up window displaying the expected risk level, categorised as either high or low, promptly following data submission. This immediate feedback mechanism is a crucial characteristic of clinical decision support technologies, as it converts backend statistics into actionable information.

The graphical user interface (GUI) was evaluated using both manual and automated data entry methods. Figure 4 illustrates the manual entry interface, which enables users to input patient information, such as age, blood pressure, and troponin levels, directly. This mode is especially beneficial in clinical settings devoid of electronic health records.

Figure 5 depicts the automated input mechanism that accommodates PDF uploads of standardised laboratory reports. The software utilises optical character recognition (OCR) and natural language processing (NLP) techniques to extract structured data from these documents. This functionality improves usability in hospital settings characterised by unstructured data.

The study proposed a conceptual framework for evaluating model decisions within a network-based paradigm, extending beyond mere classification. Each patient instance was represented as a node in a similarity graph, with the Euclidean distance between clinical characteristic vectors constituting the edges. This abstraction facilitates the mapping of classification results as node labels inside a complicated network framework.

The graph-theoretic depiction of patient data in Figure 2 establishes a fundamental framework for extrapolating to other biomedical applications, such as microbiome-host interaction networks or gene co-expression networks. In these circumstances, nodes may represent microbiological taxa or genes, and the classifier's output can be interpreted as health-state indicators situated within topological space. This strategy aligns with ongoing initiatives in computational biology to utilize graph-based reasoning and spectral learning techniques for predictive modelling in high-dimensional biological systems.

IV. DISCUSSION

The creation and assessment of the Heart Attack Prediction Software validate the efficacy of machine learning (ML) models in medical risk evaluation, particularly in the context of cardiovascular diseases. Among the implemented algorithms, K-Nearest Neighbours (KNN), Naïve Bayes (NB), and Decision Tree (DT), the Decision Tree classifier exhibited superior performance, with an accuracy of 98%. This finding aligns with prior studies, emphasising the Decision Tree's efficacy in modelling structured clinical data, owing to its ability to capture nonlinear relationships and feature interactions with minimal processing demands [1], [9].

The software offers both manual and automated (PDF-based) data entry methods, enabling real-time, user-friendly risk estimates. Its intuitive user interface guarantees accessibility for non-expert users in healthcare environments. Upon submission of patient data, the backend machine learning engine produces an interpretable and prompt forecast concerning the individual's risk of a heart attack, which is presented through a pop-up notification. This enables swift, informed decision-making by physicians and corresponds with ongoing initiatives to develop actionable clinical decision support systems.

This work possesses significant scientific value due to its potential for generalisation and theoretical modelling, in addition to its predictive function. Each patient is seen as a node in a similarity network, with relationships between patients determined by Euclidean distances in the clinical feature space. This study links the machine learning task with network medicine by conceptualising classification as a node labelling process within the graph. This expanding discipline conceptualises disease phenotypes as disturbances in network connection and modularity [2].

Figure 2 illustrates this conceptual bridge, depicting a simplified network architecture in which a new, unclassified patient node (in red) is connected to previously observed, labelled nodes (in blue). The ultimate categorization arises from the predominant class of the nearest neighbours, embodying the rationale behind the KNN algorithm within a network context. The graphic also serves as a visual bridge between flat feature-based learning and topological pattern recognition.

This abstraction facilitates the possible utilisation of the system in microbiome analysis, wherein network architectures depict microbial communities and their correlations with health or disease conditions. Recent research validates that both supervised and unsupervised machine learning algorithms can proficiently categorise microbiome profiles and reveal biological patterns in high-dimensional taxonomic data [6].

This abstraction facilitates the prospective use of the system in microbiome and genetic network analysis, where graph structures denote microbial taxa or gene modules and their relationships with health conditions. In these instances, nodes signify biological entities, while edges denote co-expression or co-occurrence patterns. Recent research employing IT-enabled WGCNA to uncover regulatory gene modules in leukaemia [7] illustrates the significance of graph-based modelling for clinical interpretation and therapeutic design. Utilising spectrum analysis on the adjacency matrix A or the Laplacian L enables researchers to get insights from eigenvalue distributions, spectral gaps, and network modularity techniques frequently employed in systems biology [4], [5].

The application of matrix functional, such as exp(A) and cosh(A), along with associated numerical approximations, facilitates the modelling of dynamic biological systems comprising millions of nodes and linkages.

Furthermore, analogous modelling techniques employing logistic equations and numerical integration approaches have demonstrated efficacy in domains such as economic forecasting, especially in elucidating inflation dynamics [8].

Despite its success, the current program implementation has its limitations. The dataset is limited in size and static; further validation on bigger, more diverse cohorts is necessary to confirm scalability and generalisability. The feature set employed was chosen empirically. Subsequent iterations of the system ought to integrate feature ranking methodologies or centrality-based metrics to enhance the predictive inputs. The model, however, utilises cross-sectional data, although cardiovascular risk is a temporally dynamic occurrence. Integrating longitudinal health records would improve the temporal resolution and precision of predictions.

This study presents a reliable and effective instrument for predicting heart attacks and establishes a basis for extending machine learning classification to network-structured biomedical challenges. The system's conceptual expansion into graph-based representations facilitates future investigations into microbiome networks, gene regulation networks, and other topologically complex topics. This study makes a significant contribution to the multidisciplinary integration of predictive analytics and systems biology by aligning machine learning outputs with graph theory and spectral analysis.

V. CONCLUSION AND FUTURE WORK

This paper details the creation and validation of a clinical decision support system, Heart Attack Prediction Software, utilising supervised machine learning models for the early identification of cardiovascular risk. The Decision Tree model demonstrated exceptional predictive performance among the evaluated classifiers, achieving 98% accuracy, thereby affirming its practical utility in primary care and emergency triage processes [1]. The software's dual-mode data input (manual and PDF-based) makes it highly adaptable to diverse health information systems, while its straightforward interface ensures accessibility for non-specialist medical staff.

This research transcends technical implementation by presenting a graph-based interpretation of machine learning predictions. Patients are represented as nodes in a similarity network, with clinical features determining the interactions between edges. This method aligns with modern trends in network medicine, where disease states are viewed as topological disturbances within biological networks [2]. The technique facilitates node-level risk classification and structural pattern detection, thus connecting machine learning with graph theory.

The report outlines numerous critical directions for advancing this effort. One entity augments the system with real-time learning capabilities, including reinforcement learning, to provide dynamic feedback and risk modification as new data is assimilated. Another approach involves incorporating patient monitoring using wearable devices (e.g., smartwatches, heart rate monitors), which would convert static risk models into dynamic health graphs by revising adjacency and attribute matrices at each temporal interval. An especially promising application is the use of this architecture to analyze microbiome data, where microbial taxa, genes, and patient samples serve as nodes within multi-layered networks. In these applications, disease prediction is based on the spectral properties of the networks, including eigenvalue distribution, centrality-weighted labelling, and spectral entropy [4], [5]. Recent toolkits combining ML and microbiome pipelines further reinforce the relevance of such approaches [11], [19].

Figure 7 exemplifies this interdisciplinary trajectory by depicting a conceptual network that amalgamates human, microbial, and genetic nodes. Each node is colour-coded according to risk or category (e.g., red for high-risk patients, green for low-risk, sky blue for microorganisms, orange for genes). At the same time, the edges represent diverse biological or clinical relationships.

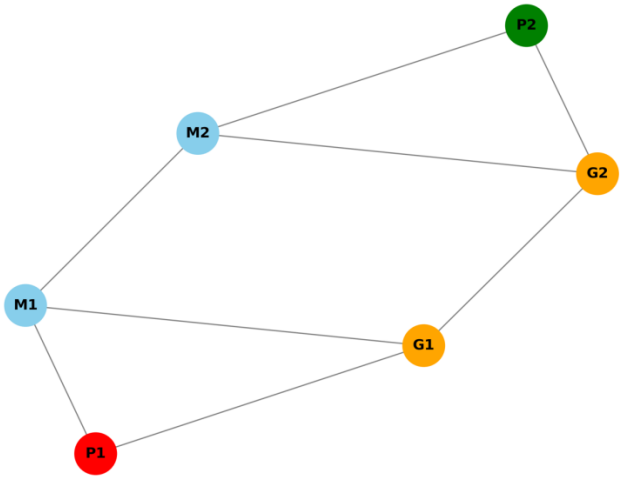

Figure 7. Vision for future graph-based integration

Nodes represent patients, microbes, or genes. Edges represent clinical similarity or biological interactions. Classifier output is embedded as node states.

Figure 7 illustrates the shift from individual classification models to topologically informed learning systems. In these systems, predictions rely not just on individual data vectors but also on the structural position and interaction context of each node within the biological network. This transition facilitates the spectral grouping of functional sub networks, temporal modelling of illness progression, and the

utilization of graph neural networks or matrix functional, such as $\exp(A)$ and $\cosh(A)$, for learning on extensive biological graphs.

This study presents a comprehensive implementation of a cardiovascular risk prediction system and a conceptual framework for analysing complex networks. It positions machine learning not just as a classification tool but also as a means to clarify biomedical complexities through network structures. This system's dual functionality, combining clinical informatics and complex systems modelling, promotes future progress in bioinformatics, network biology, and computational epidemiology. Moreover, analogous machine learning methodologies have demonstrated a favourable influence in energy management systems, facilitating intelligent control, optimisation, and sustainability using AI-driven decision frameworks [16].

REFERENCES

[1] Alshraideh, M., Alshraideh, N., Alshraideh, A., Alkayed, Y., Trabsheh, Y. A., & Alshraideh, B. (2024). Enhancing heart attack prediction with machine learning: A study at Jordan University Hospital. *Applied Computational Intelligence and Soft Computing*, 2024, Article ID 5080332. https://doi.org/10.1155/2024/5080332

[2] Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

[3] Barabási, A.-L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews Genetics*, 12(1), 56–68. https://doi.org/10.1038/nrg2918

[4] Estrada E., The structure of complex networks, Theory and Application, Oxford University Press, 2011 ISBN 978-0-19-959175-6.

[5] Gao, J., Lin, Y., & Wang, J. (2019). Spectral characterization of microbiota in health and disease. *PLoS Computational Biology*, 15(7), e1007146. https://doi.org/10.1371/journal.pcbi.1007146

[6] Kalluçi, E., Preni, B., X. Dhamo, Noka, E., Bardhi, S., Bani, K., Macchia, A., Bonetti, G., Dhuli, K., Donato, K., Bertelli, M., Zambrano, L. J. M., Janaqi, S., "A comparative study of supervised and unsupervised machine learning algorithms applied to human microbiome", Clin Ter 2024; 175 (3):98-116, https://doi.org/10.7417/ct.2024.5051, https://clinicaterapeutica.it/ojs/index.php/1/article/view/882/636

[7] Kapçiu, R., Preni, B., & Kalluçi, E. (2024a). IT-Enabled WGCNA for Critical Gene Module Mapping and Therapy Optimization: Advancing Leukemia Care. Transdisciplinary Journal of Engineering & Science, 15. https://doi.org/10.22545/2024/00252

[8] Kapçiu, R., Preni, B., Kalluçi, E., & Kosova, R. (2024b). MODELING INFLATION DYNAMICS USING THE LOGISTIC MODEL: INSIGHTS AND FINDINGS. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi|JIITUJ|*, *8*(1), 364–378. https://doi.org/10.22437/jiituj.v8i1.32605

[9] Kapil, A. R. (2022). Decision tree algorithm in machine learning: Advantages, disadvantages, and limitations. *AnalytixLabs*. https://www.analytixlabs.co.in/blog/decision-tree-algorithm

[10] Liu, Y. Y., Slotine, J. J., & Barabási, A. L. (2020). Controllability of complex networks. *Nature*, 473(7346), 167–173. https://doi.org/10.1038/nature10011

[11] Marcos-Zambrano, L. J., López-Molina, V. M., Bakir-Gungor, B., Frohme, M., Karaduzovic-Hadziabdic, K., Klammsteiner, T., Ibrahimi, E., Lahti, L., Loncar-Turukalo, T., Dhamo, X., Simeon, A., Nechyporenko, A., Pio, G., Przymus, P., Sampri, A., Trajkovik, V., Lacruz-Pleguezuelos, B., Aasmets, O., Araujo, R., . . . De Santa Pau, E. C. (2023). A toolbox of machine learning software to support microbiome analysis. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1250806

[12] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

[13] Rashid, T. A., & Hassan, B. (2022). Heart attack dataset. *Mendeley Data*, V1. https://doi.org/10.17632/wmhctcrt5v.1

[14] Saravanan, T., Raghavaraju, D., Bhaskar, M. V., & Vasundra, S. (2023). Prediction of heart disease using machine learning and hybrid methods. In *Proceedings of the 2023 1st International Conference on Optimization Techniques for Learning*. https://doi.org/10.1109/ICOTL59758.2023.10435033

[15] Sherrell, Z. (2021, October 22). What to know about blood tests to diagnose a heart attack. *Medical News Today*. https://www.medicalnewstoday.com/articles/blood-tests-for-heart-attack

[16] Tabaku, E., Vyshka, E., Kapçiu, R., Shehi, A., & Smajli, E. (2025). UTILIZING ARTIFICIAL INTELLIGENCE IN ENERGY MANAGEMENT SYSTEMS TO IMPROVE CARBON EMISSION REDUCTION AND SUSTAINABILITY. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi|JIITUJ|*, *9*(1), 393–405. https://doi.org/10.22437/jiituj.v9i1.38665

[17] Venkatesh, D., Saravanan, T., Raghavaraju, D., Bhaskar, M. V., & Vasundra, S. (2023). Prediction of heart disease using machine learning and hybrid methods. *2023 1st International Conference on Optimization Techniques for Learning*. https://doi.org/10.1109/ICOTL59758.2023.10435033

[18] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442. https://doi.org/10.1038/30918
[19] World Health Organization. (2023). *Cardiovascular diseases (CVDs)*. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)